

---

**Improved methods for the formation and stabilization of R-loops**

---

---

David B.Kaback, Lynne M.Angerer\* and Norman Davidson

---

---

Department of Chemistry, California Institute of Technology, Pasadena, CA 91125, USA

---

---

Received 23 January 1979

---

ABSTRACT

*Improved methods for the formation and stabilization of R-loops for visualization in the electron microscope are presented. The two complementary strands of a duplex DNA are photochemically crosslinked once every 1 to 3 kb using 4, 5', 8 trimethylpsoralen. R-loops are then formed by incubation with RNA in 70% formamide at a temperature above the DNA melting temperature. Finally, the R-loops are stabilized by modifying the free single strand of DNA with glyoxal, thus minimizing the displacement of the hybridized RNA by branch migration. In this manner R-loops can be formed and visualized at a high frequency irrespective of the base composition of the nucleic acid of interest.*

INTRODUCTION

Under appropriate conditions RNA will hybridize with double stranded DNA displacing the non-complementary DNA strand. The structure that forms has been termed an R-loop, and is easily recognized in the electron microscope as a bubble in the otherwise duplex DNA structure. The bubble is comprised of a double stranded DNA:RNA hybrid half and an single stranded DNA half (1, 2). R-loops have been useful for studying the organization of transcribed sequences in viral and cloned recombinant eucaryotic DNA's (3,4,5,).

Rapid and efficient formation of R-loops occurs in solvents in which RNA:DNA hybrids are more stable than DNA duplexes and at temperatures very close to or above the DNA:DNA strand separation temperature ( $T_{ss}$ ) of the sequence that hybridizes to the added RNA (1). To decrease the probability of total strand separation, it is advantageous if the DNA sequence which is going to form an R-loop is flanked by sequences with a higher  $T_{ss}$  that remain hydrogen bonded during hybridization. Accordingly, formation of R-loops with G + C rich RNA is difficult because the two DNA strands may irreversibly dissociate before an R-loop forms.

A further difficulty may be encountered when trying to keep high molecular weight chromosomal and viral DNA molecules intact during R-loop formation.

These DNA molecules are often nicked. Under conditions favoring R-loop formation the DNA duplex between two nicks on opposite strands may dissociate resulting in a reduction of the duplex DNA length.

These problems can be overcome and R-loops readily formed if the two complementary DNA strands are crosslinked infrequently with 4, 5', 8-trimethylpsoralen (trioxsalen). Trioxsalen intercalates into duplex DNA and forms covalent crosslinks between adjacent pyrimidine residues on complementary strands when irradiated with long wavelength ultraviolet light (6). The number of crosslinks can be controlled by varying the dose of ultraviolet irradiation (6). In this paper we show that DNA infrequently crosslinked with trioxsalen can be used to form R-loops with high efficiency.

Once formed, R-loops are not completely stable if the molecules are transferred to certain altered solvent conditions. For example, we found the majority of yeast mRNA:DNA and rRNA:DNA R-loops formed in 70% formamide were only partially stable in 50% formamide, a solvent that gives better contrast in the electron microscope. In 50% formamide R-loops are less stable thermodynamically and part or all of the R-loop may be displaced by branch migration of the unhybridized DNA strand. We have devised a method which stabilizes R-loops by preventing branch migration. R-loops are treated with glyoxal under conditions where only the single stranded DNA arm is modified, thus inhibiting further base pairing while preexisting duplex DNA is not visibly affected. This method has made it possible to stabilize R-loops for electron microscopic visualization in several systems for which R-loop instability has been a problem.

### MATERIALS AND METHODS

Reagents and their abbreviations are as follows: TE--0.01 M Tris, 0.001 M EDTA, pH 7.4. NPE--0.5 M NaCl, 0.1 M PIPES, 0.01 M EDTA, pH 7.2. PE--0.1 M PIPES, 0.01 M EDTA, pH 7.2. Hybridization Buffer (R-loop buffer)--70% formamide (v/v) recrystallized as described (7), 0.4 M NaCl, 0.1 M PIPES, 0.01 M EDTA, pH 7.2 (1). PLB II--0.1 M NaCl, 0.02 M EDTA, pH 7.0 (8). DM--2.0% (w/v) Na lauroyl sarcosinate. 3.0% (w/v) Na deoxycholate, 5.0% (w/v) sodium dodecyl sulfate, 0.02 M EDTA, 0.01 M Tris, pH 8.0 (8).

Nucleic Acid Isolation. High molecular weight (100-300 kilobases (kb)) yeast DNA was isolated by a modification of the spheroplast lysis method (8) from log phase cells grown on YEPD (1.0% (w/v) yeast extract, 2.0% (w/v) peptone and 2.0% (w/v) glucose). Spheroplasts from  $1.5 \times 10^9$  -  $3 \times 10^9$  cells were prepared according to Blamire et al. (8), suspended in 2 ml PLB II,

layered on top of the complex gradient pictured in Fig. 1 in an SW 25.2 centrifuge tube and centrifuged at 18,000 rpm for 8 hr at 15°C. The DNA banded in the middle of the CsCl layers. The top 35 ml of the gradients were discarded and the remaining solutions carefully mixed and adjusted to a density of 1.7 gm/cc ( $\eta = 1.400 \pm 0.005$ ) with saturated CsCl solution. Care was taken to minimize any agitation beyond what was necessary to obtain uniform solutions. The gradients were centrifuged in a 60 Ti rotor at 36,000 rpm for 48-60 hr at 20° and fractionated by puncturing the bottom of the tube with a 1" long 12 gauge needle and pumping at a rate of approximately 1 ml per min. Emerging drops were collected by direct contact with the bottom of the tube to minimize mechanical shear forces. Finally the fractions containing DNA

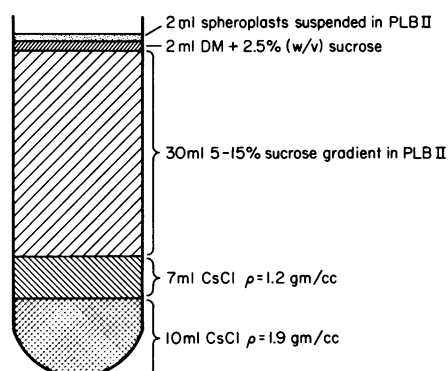


Figure 1. Complex gradient for isolation of high molecular weight DNA. Just prior to running, the two CsCl solutions were freshly underlayered and the DM and spheroplast solutions freshly overlaid on the sucrose gradient.

were dialyzed against TE. Up to 50  $\mu$ g of DNA were obtained per gradient. It is necessary to use both the CsCl layers and the sucrose gradient in order to remove cellular debris and a nondialyzable contaminant which acts as a surfactant that interferes with the spreading of DNA for electron microscopy.

The entire cellular DNA content was obtained by this preparation as judged by the presence of mitochondrial and 2  $\mu$  circular DNA, which should be the slowest sedimenting DNA in the first step (data not shown). DNA isolated in this manner was at least 100-300 kb in length.

(It may have been larger since our only criteria for judging the molecular weight was electron microscopy after Kleinschmidt spreading, which may shear the DNA.) The DNA was relatively free of nicks compared to DNA isolated by conventional methods which utilize enzymes (data not shown). The ratio of  $A_{260}/A_{280}$  always approached 2.0 indicating it was relatively free from protein.

pJHC11 plasmid DNA (9) was isolated from chloramphenicol treated *E. coli* HB101 according to the method of Wensink et al. under P2 containment conditions (10). pJHC11 was digested with restriction endonuclease BamHI (N.E. Biolabs), phenol extracted and dialyzed against TE.

18S and 25S rRNA were isolated as previously described (11). poly(A)<sup>+</sup>

RNA (the gift of F. Gibson) was isolated as described (12) from log phase YEPD grown cells.

Crosslinking with Trioxsalen. DNA was suspended in 0.01 M  $\text{NaHPO}_4$ , 0.001 M EDTA pH 6.9 at a concentration of 50  $\mu\text{g/ml}$ . Trioxsalen (50  $\mu\text{g/ml}$  in absolute ethanol) was added to a concentration of 1  $\mu\text{g/ml}$  and the solution irradiated for varying times (1-5 min) at a distance of 30 cm from a Black-Ray long wave length ultraviolet lamp (output at 30 cm = 45  $\text{ergs/mm}^2\text{sec}$ ) (Ultraviolet Products, Inc., San Gabriel, CA). Care was taken to get a uniform distribution of trioxsalen dissolved in the sample. The sample was gently mixed 2-3 times during the irradiation. As the number of crosslinks varies for different types of DNA (unpublished observation), irradiation was generally done for 3 different times and the frequency of crosslinks checked.

Determination of the Crosslinking Frequency. Aliquots of crosslinked DNA were spread in 50% formamide after complete denaturation with 1.0 M glyoxal as described (13). The number of single stranded crossovers was counted and divided by the length of the DNA measured. The lengths of the two strands between a crossover had to be approximately equal for the crossover to be counted as a crosslink.

R-loop Formation. R-loops were formed by incubation of RNA with DNA in hybridization buffer. Incubation temperatures and times for particular experiments are noted in the text.

Glyoxal Fixation of R-loops. The R-loop mixture was quickly chilled to 0°, 40% glyoxal (Matheson, Coleman and Bell) was added to a final concentration of 1.0 M (1/7 volume), and the solution was incubated 2 hr at 12-13°C. Care was taken to prevent any warming of the R-loop solution when it was in the presence of 1.0 M glyoxal. R-loops were either spread directly or dialyzed against TE or NPE at 4°.

Separation of Unhybridized RNA from R-looped DNA. 50  $\mu\text{l}$  of the poly(A)<sup>+</sup> RNA R-loop solution was fixed with glyoxal, dialyzed against NPE and chromatographed on a 3 ml Sepharose 2-B column equilibrated with NPE. 0.2 ml fractions were collected and an aliquot of each fraction spread in 50% formamide and examined in the electron microscope. All DNA molecules including those containing R-loops were excluded while most of the unhybridized mRNA was included.

Spreading for Electron Microscopy. R-loop preparations in 50% formamide were spread onto a 15% formamide hypophase as described (14). Spreading

from 70% formamide on a 15% formamide hypophase was carried out as described by Chow et al. (15) for 45% formamide spreads. All components of the spreading solutions, except the R-looped DNA were mixed. The R-loop mixture was added and this spreading solution immediately pipetted onto the hypophase. The limited exposure of R-loops to spreading conditions minimizes displacement of the RNA by branch migration (T. Broker and L. Chow, personal communication). Films were stained with uranyl acetate and rotary shadowed with Pt/Pd. Electron microscopy was done on a Philips 300 instrument at varying magnifications using  $\phi$ X 174 RFII DNA (5386 base pairs) (16) as a double stranded length standard.

## RESULTS

rDNA R-loop Map of pJHCII. Initial tests on various R-loop procedures were carried out on the plasmid pJHCII which contains yeast rDNA fused to the bacterial drug resistance vector RSF2124 (9). After conversion to a full length linear form by digestion with Bam HI, pJHCII DNA was incubated with 18S rRNA or 18S and 25S rRNA in hybridization buffer for 3 hr at 50° and 52° respectively. The DNA was spread in both 50% and 70% formamide and observed in the electron microscope. Figure 2 shows the R-loop maps generated. When 18S rRNA alone is hybridized, 93% of the 18S rRNA genes contained a recognizable R-loop (Fig. 2c). Hybridizing both 18S and 25S rRNA, greater than 95% of all the 18S and 25S rRNA genes contained R-loops (Figs. 2a and b). An additional small R-loop from the 3' end of the 25S rRNA was reproducibly present only when the R-loops were spread in 70% formamide. A typical micrograph of Bam HI cut pJHCII containing 18S and 25S rRNA R-loops is shown in Fig. 3. We conclude that pJHCII contains almost two tandem repeats of the rRNA genes, including the two intact 18S rRNA genes, an intact 25S rRNA gene and a fragment of a second 25S rRNA gene. These results agree with both the restriction map of pJHCII (9) and the results obtained by hybridizing RNA to the purified rDNA EcoRI restriction fragments (17,18,19).

Effects of Crosslinking on R-loop Structure. Bam HI digested pJHCII, that had been trioxalen crosslinked once per 740 base pairs, was incubated with 18S rRNA in hybridization buffer for 2-3 hr at 50°C and spread from 50% formamide. 84% of the 18S rRNA genes contained a detectable R-loop. The 18S RNA R-loop map using the crosslinked DNA (Fig. 2d) was almost identical to the control containing unmodified DNA spread in 50% formamide (2c) with the exception that the average R-loop is 15% shorter in the crosslinked sample.

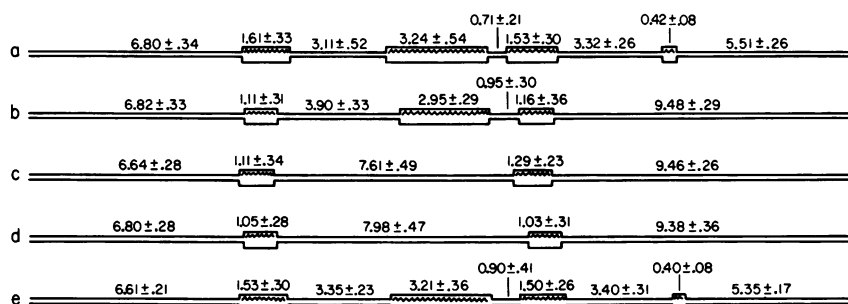


Figure 2. Maps of 18S and 25S rRNA R-loops on Bam-HI digested pJCH11 DNA prepared under various conditions. R-loops were formed and prepared for electron microscopy as described in Materials and Methods. The sample size for each map was 40-60 molecules. The lengths of each of the segments in all the maps were normalized to a total length of 26.20 kb, the mean combined length of a-e. a. 18S and 25S rRNA R-loops spread from 70% formamide, b. 18S and 25S rRNA R-loops spread from 50% formamide, c. 18S rRNA R-loops spread from 50% formamide, d. 18S rRNA R-loops on trioxsalen crosslinked (once per .74 kb) DNA spread from 50% formamide, e. glyoxal fixed 18S and 25S rRNA R-loops spread in 50% formamide. Errors given are the standard deviations. RNA is represented by the jagged line.

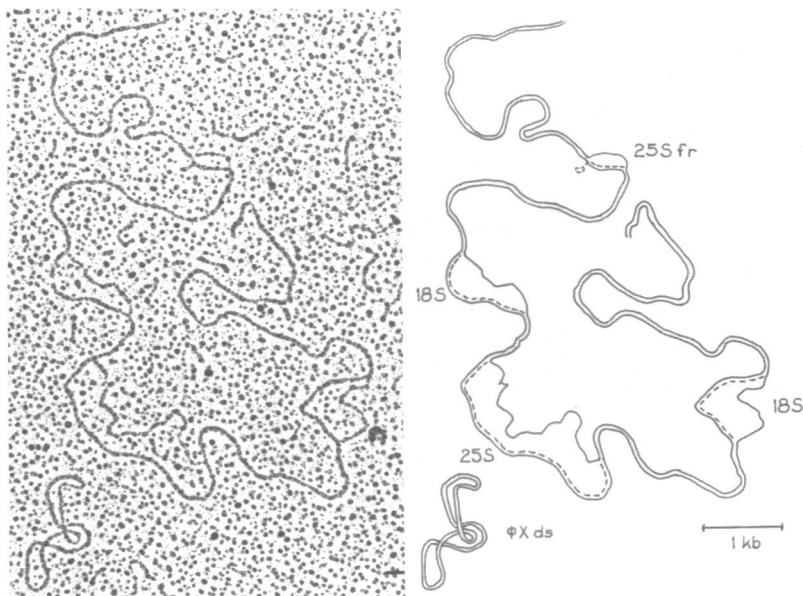


Figure 3. Electron micrograph of 18S and 25S rRNA R-loops on Bam-HI digested pJCH11 DNA spread from 70% formamide as described in Materials and Methods. The 25S rRNA gene fragment (25S fr) shows a single stranded tail, presumably made up of the 5' end of the rRNA molecules. RNA is represented as a dashed line.

R-loops interpreted to contain several crosslinks in a single gene were sometimes observed (Fig. 4a). Structures were also observed where a single gene contained two R-loops spaced by a region that was either duplex or triplex (Fig. 4b). These structures may be indicative of too many crosslinks interfering with either the formation or visualization of the R-loops in the center of the gene. Accordingly, the noticeably smaller R-loops may also be caused by too many crosslinks at the ends of the gene. Experiments reported in later sections on yeast chromosomal DNA show that when the cross-linking frequency is reduced to about once per 3 kb, the length of the R-loops is indistinguishable from that with uncrosslinked rDNA. In summary, these initial experiments show that R-loops can be formed using trioxsalen cross-linked DNA.

Formation of R-loops to Completely Denatured DNA. As stated previously, it is difficult to form R-loops when the RNA is G + C rich since the DNA

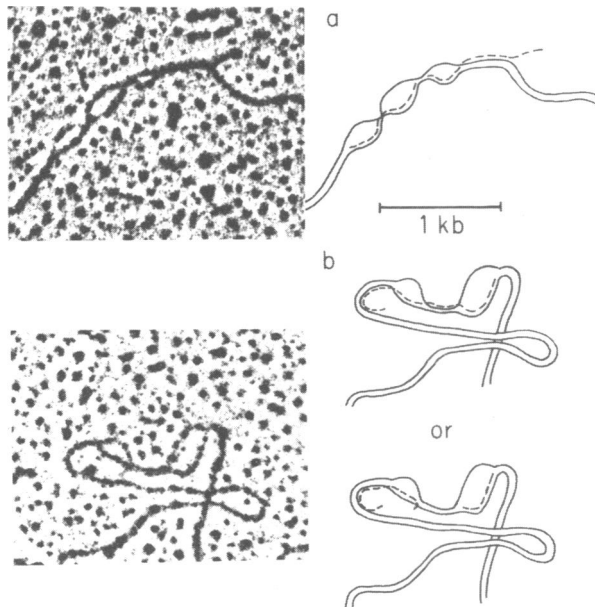


Figure 4. Examples of anomalous 18S rRNA R-loops on crosslinked (once per 0.74 kb) pJHC11 DNA spread in 50% formamide. a. In addition to the two apparent crosslinks in the center of the gene there appears to be what we interpret as several crosslinks at the end of the gene giving a collapsed triplex DNA:DNA:RNA structure. b. An R-loop that has collapsed around the middle of the gene. It is not clear if the center region is a duplex DNA or a triplex DNA:DNA:RNA structure.

strands can irreversibly separate during the hybridization. To test if crosslinking can be used to prevent irreversible denaturation of the DNA and allow R-loop formation, we hybridized 18S rRNA to crosslinked Bam HI digested pJHC11 DNA under conditions where the DNA is completely denatured. A thermal denaturation curve, monitored optically, of Bam HI cut pJHC11 in R-loop hybridization buffer showed that the midpoint of the thermal denaturation curve ( $T_m$ ) was 48.5 and the DNA was completely denatured at 58° (not shown). In 70% formamide the  $T_m$  of DNA:RNA hybrids is 15-25° higher than the  $T_m$  of DNA duplexes (7). Therefore, at temperatures above 58° rRNA:DNA hybrids are stable (estimated  $T_m > 70^\circ$ ) while the plasmid DNA duplex is dissociated unless held together by trioxsalen crosslinks.

18S rRNA was hybridized to unmodified and crosslinked (once per 740 base pairs) pJHC11 DNA in hybridization buffer for 4 hrs at 63°C. The two samples were spread in 50% formamide and the percentage of R-loops determined by electron microscopy. The unmodified DNA was, as expected, almost completely single stranded. Furthermore, most of the single strands were shorter than full length. A few (<10% of predicted) R-loops were seen. These were probably due to a small amount of renaturation of complementary DNA strands which occurred when the reaction mixture was cooled to room temperature just prior to spreading. In contrast, the majority of the molecules in the crosslinked sample were full length duplexes containing 89% of the predicted number of 18S R-loops, indicating R-loops can form at higher temperatures when DNA is completely single stranded.

Formation of R-loops with Large Chromosomal DNA. rDNA comprises 6-8% of yeast DNA (20,21) so that it is easy to locate rDNA R-loops in spreads of unfractionated DNA. When 18S rRNA was hybridized in R-loop buffer for 2-3 hrs to 100-300 kb yeast chromosomal DNA at 50°, the largest rDNA molecules observed were 75 kb long and contained no more than 8 regularly spaced R-loops. The majority of the 18S RNA R-loops were present on significantly shorter molecules (27-45 kb) containing only 3-5 R-loops. In contrast, when long yeast chromosomal DNA (100-300 kb) was crosslinked on the average once per 2.75 kb and hybridized with 18S rRNA under the same conditions, very long strands of DNA with lengths up to 236 kb containing 26 regularly spaced 18S rRNA genes were found. A molecule containing 15 repeats is shown in Fig. 5. Only 25-35% of all the rDNA was present on molecules containing fewer than 6 repeats (55 kb). The average length of the 18S rRNA was  $1.22 \pm .21$  kb which was comparable to the length found in the 50% formamide spreads of 18S rRNA R-loops on unmodified pJHC11 molecules. The





Figure 5. A molecule of yeast rDNA containing 15 approximately evenly spaced 18S rRNA-loops. The DNA:RNA hybrid strand is represented as a double line and the unpaired DNA strand is represented as a dashed line.

average length of the repeating R-loop plus spacer unit was  $9.34 \pm .18$  kb. Restriction enzyme analysis of the yeast rDNA gives a repeat length of 9.1 kb (18,19) in good agreement with our measurements.

The frequency with which we saw spacings between 18S R-loops of  $18.7 \pm 2.0$  kb rather than 9.3 kb was about 5%. This is expected on the basis of the observed 94% saturation of 18S rRNA R-loop sites in pJHC11 DNA treated under similar conditions. This assumes that all yeast rDNA consists of 9.3 kb tandem repeats with no major length heterogeneity.

R-loops were also formed with both 18S and 25S rRNA at 53° using crosslinked (once per 2.7 kb) unfractionated yeast DNA. Molecules containing 8 tandem repeats of 18S and 25S rRNA genes were observed, but no attempt was made to get longer DNA on the grids. The tandem repeat lengths for the genes and spacer were in complete agreement with those obtained with 18S rRNA alone. In addition, the relative positions of the 18S and 25S rRNA R-loops were in agreement with those observed for pJHC11.

In summary, the experiments described in this section show that infrequent trioxsalen crosslinking is effective for keeping large duplex molecules relatively intact during the formation of R-loops.

Stabilization of rRNA R-loops with Glyoxal. In the previous section we have dealt with the problem of keeping the DNA intact during R-loop formation. An additional problem encountered is the instability of R-loops under certain conditions. The length of the 18S rRNA is  $1.80 \pm .20$  kb as determined by  $\text{CH}_3\text{HgOH}$  gel electrophoresis (unpublished observation) and a variety of other techniques (18,22) while the length of 18S yeast rRNA R-loops formed in 70% formamide and spread from 50% formamide was only 1.11-1.29 kb (Fig. 2c). Furthermore, none of the R-loops spread in the 50% formamide were full length. In contrast, 18S rRNA R-loops spread from 70% formamide were  $1.57 \pm .30$  kb long. 50% of these R-loops were full length and many of the rest contained a small tail of unhybridized RNA. The length of this tail plus the length of the R-loops was approximately 1.8 kb. Since DNA:RNA hybrids are thought to have almost the same (4% shorter) contour length per kb in the electron microscope as DNA duplexes (18), the shortening of the R-loops spread from 50% formamide is probably the result of branch migration as opposed to an artifact of the spreading condition.

We found that treatment of DNA containing R-loops with 1.0 M glyoxal in 70% formamide at low temperature is useful for stabilizing R-loops against displacement by branch migration. Glyoxal binds covalently to G residues on single strand nucleic acids, and destroys their ability to form base pairs

(23). Under sufficiently gentle conditions, duplex DNA is not visibly denatured (24).

18S and 25S rRNA R-loops on pJHC11 were formed and treated with 1.0 M glyoxal at 12-13° for 2-3 hr as described in Materials and Methods. The DNA was either directly spread from 50% or first dialyzed against TE and then spread from 50% formamide. Dialysis made no quantitative difference; however, the quality of the electron microscopic image was improved. In either case, as shown in Fig. 2e, R-loops fixed with glyoxal are significantly larger than the unfixed R-loops spread from 50% formamide and are about the same size as the R-loops spread from 70% formamide. 36% of the glyoxal fixed R-loops were within 10% of the expected full length of 1.8 kb.

In addition to the shortening of R-loops caused by spreading from 50% formamide, we have observed complete loss of some R-loops under these conditions. As shown in Fig. 2, pJHC11 contains a  $0.42 \pm .08$  kb segment from the 3' end (25) of a 25S rRNA gene at one of the yeast-vector junctions. This R-loop is almost always present when the molecules are either spread directly after the hybridization reaction from 70% formamide or fixed with glyoxal and then spread in 50% formamide. Furthermore, once fixed with glyoxal they are completely stable when dialyzed against TE. However, in the absence of glyoxal fixation the 25S rRNA fragment R-loop spread from 50% formamide is either partially or entirely lost in 66% of the molecules observed. The various structures and their observed frequencies when spread under different conditions are shown in Table I. These results show that glyoxal modification at low temperatures permits spreading in 50% formamide with equal or slightly greater R-loop stability than spreading from 70% formamide.

R-loop Formation with Yeast poly(A)<sup>+</sup> RNA. To test these R-loop procedures on a large heterogeneous sample of coding sequences, total yeast poly(A)<sup>+</sup> RNA was hybridized under saturating conditions to total yeast DNA crosslinked once every 2.2 kb with trioxsalen. In order to insure complete DNA strand separation, R-loop incubations were carried out in hybridization buffer at 52°C. This temperature is 1-3° higher than the irreversible  $T_{ss}$  of total yeast DNA in hybridization buffer as estimated from melting profiles of total yeast DNA in aqueous buffers with a cation concentration of 0.2 M (J. Cramer, personal communication). 1 mg/ml poly(A)<sup>+</sup> RNA was incubated with 10 µg/ml yeast DNA for 105 hr ( $R_o t = 1.16 \times 10^3$  M sec). This value is 50 times greater than the  $R_o t$  used to reach saturation in an mRNA driven hybridization to denatured single copy DNA (12). Since the rate of hybridization is

TABLE 1 EFFECT OF DIFFERENT SPREADING CONDITIONS ON THE STABILITY OF THE 25S rRNA GENE FRAGMENT IN pJHC11.

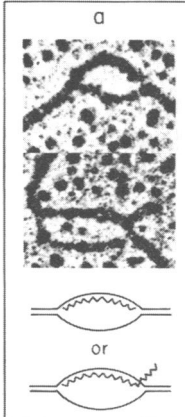

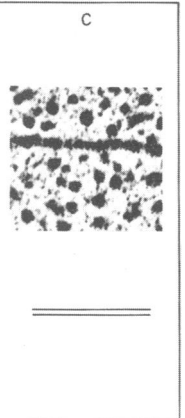
SPREADING CONDITIONS	NUMBER OF OBSERVED STRUCTURES		
	a	b	c
			
70% FORMAMIDE	47	2	1
50% FORMAMIDE	33	29	37
GLYOXAL TREATED 50% FORMAMIDE	47	0	1
GLYOXAL TREATED DIALYZED AGAINST TE, 50% FORMAMIDE	52	0	0

Table 1. Effect of Different Spreading Conditions on the Stability of the 25S rRNA Gene Fragment in pJHC11. a. plasmids containing a discernible R-loop. In some molecules a tail presumed to be the unhybridized 3 kb of the 25S RNA was observed. When no tail was present it was presumed to be due to either collapse of the RNA in the spread or the hybridization of degraded RNA. b. plasmids containing a single stranded RNA tail attached to DNA in the correct region of the plasmid. These were thought to be due to partial displacement of the R-loop by branch migration. c. plasmids containing no structures in the appropriate region presumed to be due to complete displacement of the R-loop by branch migration. A small proportion (<5%) may be due to the failure to form R-loops.

slower in high formamide solvents (7) and the rate of R-loop formation is temperature dependent (1), this high  $R_{ot}$  value was chosen to insure saturation under R-loop hybridization conditions (We have subsequently observed that this  $R_{ot}$  is at least 3 times greater than necessary to reach saturation.) DNA containing R-loops was glyoxal treated as described in Materials and Methods. It was either spread from 50% formamide or first dialyzed against NPE and passed over a Sepharose 2B column before spreading. Controls were not treated with glyoxal and were immediately spread from 50% or 70% formamide without dialysis. The data in Table 2 show that the number of R-loops per

TABLE 2  
Stabilization of poly(A)<sup>+</sup> RNA R-loops after Treatment with Glyoxal

Experimental Conditions	Total kb DNA traced	kb DNA in R-loops	% of total DNA in R-loops	number R-loops	Frequency of R-loops per unit length of DNA <sup>1</sup>	R-loops per yeast genome <sup>2</sup>	Average R-loop length(kb)	Average size double stranded DNA between adjacent R-loops
Spread from 50% formamide <sup>4</sup>	351.1	23.1	6.6	55	0.16	2424	0.51±0.22 <sup>3</sup>	4.65±4.60
Spread from 70% formamide <sup>4</sup>	544.0	51.5	9.5	101	0.19	2813	0.50±0.30	4.16±4.57
glyoxal fixed and spread from 50% formamide	399.5	130.7	32.6	124	0.31	4698	1.06±0.33	2.28±2.29
glyoxal fixed, dialyzed, gel filtered and stored 30 days, 4°C and spread from 50% formamide	525.0	152.0	28.9	174	0.33	5000	0.87±0.39	1.84±2.25

Yeast poly(A)<sup>+</sup> RNA hybridized to crosslinked total yeast DNA under R-loop conditions was quickly chilled to 0°C and treated as described above. A small aliquot was spread at a sufficient dilution so that the unhybridized RNA did not interfere with the R-looped DNA.<sup>5</sup> All the DNA on a portion of the grids was photographed and the number and lengths of the R-loops determined.

<sup>1</sup> Frequency is given as the number of R-loops divided by the total kb of DNA traced.

<sup>2</sup> R-loops per genome is calculated as the frequency multiplied by the chromosomal genome size, taken as 15,152 (26).

<sup>3</sup> Errors are given as the standard deviation.

<sup>4</sup> R-loops were also spread directly from the hybridization reaction without prior chilling. No quantitative changes were observed in the chilled vs. unchilled samples and the data presented is the combined total of both.

<sup>5</sup> In the experiment presented in the bottom row of the table, RNA was removed by gel filtration so a larger aliquot of DNA could be spread.

unit length of DNA, the fraction of the DNA observed in R-loops and the average R-loop length were all much greater for the glyoxal treated samples than for the untreated controls even when the latter was spread from 70% formamide. In addition, we noted that the frequency of R-loops remained constant for a glyoxal treated sample after storage in the cold for 30 days. The average size of these R-loops appear 18% shorter than unstored glyoxal fixed R-loops, indicating that a small amount of branch migration may have occurred during storage. However, the constant frequency indicates that few, if any, R-loops are completely displaced. A typical stretch of DNA containing poly(A)<sup>+</sup> RNA R-loops fixed with glyoxal and spread from 50% formamide is shown in Fig. 6.

Studies by several other methods indicate that 3000 to 5000 genes coded for by 30-40% of the yeast genome are expressed as poly(A)<sup>+</sup> RNA (12). The percentage of DNA in R-loops observed in the glyoxal treated sample approached this value. Multiplying the number of R-loops/kb of DNA traced by the

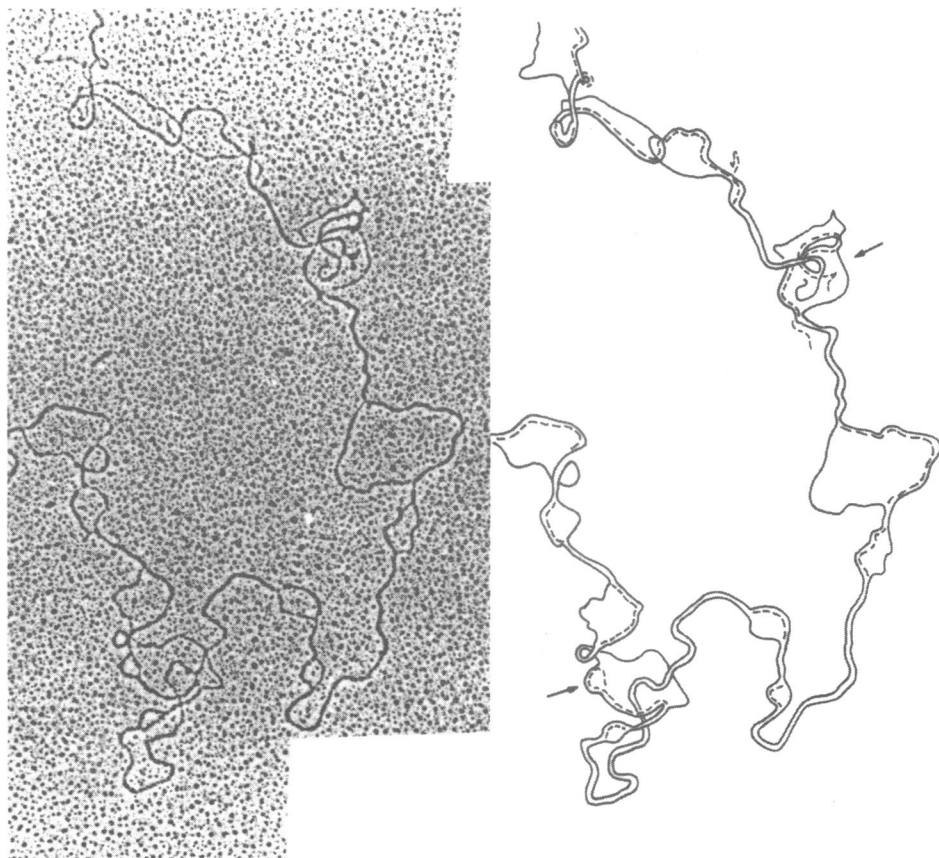


Figure 6. Yeast DNA containing poly(A)<sup>+</sup> RNA R-loops that was glyoxal fixed, dialyzed, gel filtered on Sepharose 2B and spread in 50% formamide as described in Materials and Methods. 38% of the duplex DNA length of this molecule is covered with R-loops. The arrows denote regions of R-loop containing DNA which are difficult to interpret.

number of kb in the genome (15,152 kb) (26) we estimate that there are approximately 5000 expressed transcripts in good agreement with values quoted above. A more complete description of the distribution of transcribed regions in the yeast genome will be presented elsewhere.

Effect of Glyoxal Treatment on Duplex DNA. Several bubbles less than 200 base pairs long were seen in the R-loop experiments with poly(A)<sup>+</sup> RNA. In these cases it was often difficult to discern the single strand:double strand structure of the R-loops. Although the relative number of these

structures was negligible compared to the total number of R-loops observed the possibility that these bubbles were not R-loops, but glyoxal induced single stranded DNA denaturation loops remained. To test this, duplex yeast DNA was treated with glyoxal for 2 hr and 24 hr at 12°C in hybridization buffer and spread from 50% formamide. Several small denaturation bubbles were observed which in a hybridization experiment could be confused with R-loop structures. The results show (Table 3) that only 0.06% and 0.3% of the DNA is denatured into short bubbles by the 2 hr and 24 hr glyoxal treatments, respectively. In the case of the 2 hr treatment, which appears adequate for stabilizing R-loops, the number of bubbles could account for no more than 1-2% of the total number of R-loops observed. Therefore, as a result of these observations small bubbles were ignored unless a clear R-loop structure was evident.

Formation of R-loops to Cloned Eucaryotic Genes. A standard procedure was adopted which should be applicable to most cloned eucaryotic genes. Chicken  $\beta$  globin chromosomal DNA cloned in  $\lambda$ Charon 4A (Ch4A $\beta$ gl from J. Dodgson and J.D. Engel) was crosslinked as described in Material and Methods and R-loops formed by incubation at 56° for 24 hrs with 5  $\mu$ g/ml of 7 day embryo reticulocyte 10S RNA. This RNA is a mixture of 80% adult  $\beta$  and 20% embryonic  $\beta$  globin sequences (27). The R-loops were fixed with glyoxal and spread in 50% formamide (Fig. 7). 80-100% of the molecules observed contained at least one R-loop while approximately 1/3 contained a second R-loop located  $2.41 \pm 0.15$  kb from the first. The major R-loop is the adult  $\beta$  globin gene and the minor R-loop (not shown) corresponds to the embryonic sequence (J.

TABLE 3  
Effect of Glyoxal Treatment on Duplex DNA

Length of glyoxal treatment	kb of DNA observed		% of DNA visibly single stranded	Number of bubbles observed	Estimated glyoxal induced bubbles per genome <sup>1</sup>
	Duplex	Single stranded			
2 hr.	500-1000	0.45	0.03	3	45-90
24 hr	260	0.80	0.31	8	466

Duplex DNA prepared as described in Materials and Methods, suspended in hybridization buffer was treated with 1M glyoxal for the times noted above and spread from 50% formamide. A portion of the grids was randomly photographed and the number and lengths of any bubbles measured.

$$^1\text{bubbles per genome} = \frac{\text{number of observed bubbles} \times 15,152 \text{ kb/yeast genome (26)}}{\text{estimated or measured total kb DNA}}$$

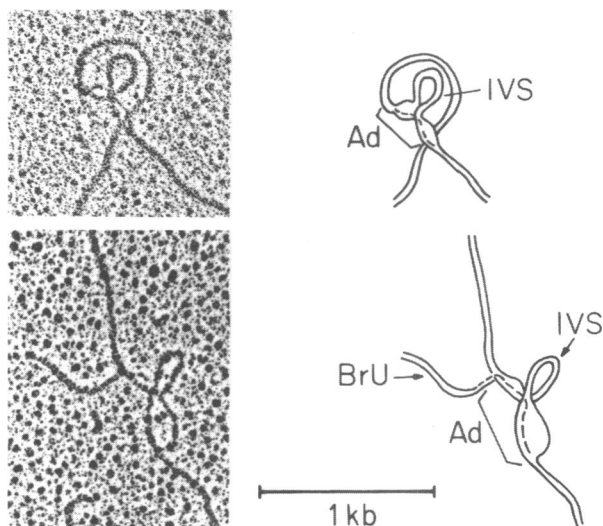


Figure 7. Adult (Ad)  $\beta$ -globin RNA R-loops on CH4A $\beta$ G1 DNA. 7 day reticulocyte 10S RNA was hybridized to crosslinked (once per 4 kb) DNA in hybridization buffer as described in the text. The free 3' poly(A) tail of the adult  $\beta$ -globin R-loop in the lower picture is labeled with poly(BrdU) tailed pMB9 restriction fragment (28).

Dodgson, J.D. Engel and J. Strommer, personal communication). Both genes contain an  $810 \pm 100$  base pair intervening sequence. Finally, since the DNA was trioxsalen crosslinked a majority of the molecules were intact after the R-loop hybridization; this was in contrast to an uncrosslinked sample of DNA.

#### DISCUSSION

Several improvements which simplify formation and visualization of R-loops in the electron microscope have been devised. The DNA is crosslinked once per 1.5 to 4 kb with trioxsalen which prevents irreversible strand separation and enables the formation of R-loops at temperatures where the DNA is completely denatured. Once formed, the R-loops are treated with glyoxal in the cold which stabilizes the structure.

There was no apparent effect on the formation of rRNA R-loops when the DNA was crosslinked with fewer than an average of one diadduct per 1.7 kb. The presence of structures such as shown in Figs. 4a and 4b suggests that the RNA just winds itself around the trioxsalen crosslink. However, if too many crosslinks are present in a single gene some R-loops may appear to be



too short or discontinuous (Fig. 4b). It is not clear whether the R-loop is not forming or if the large number of crosslinks prevents the resolution of the bipartite R-loop structure.

An appreciable fraction of covalently bound trioxsalen molecules are monoadducts bound only to a single strand (29). Although not investigated, it seems unlikely that the mild trioxsalen treatment used in these studies could cause a significant destabilization of the double helix. Duplexes viewed in the electron microscope appeared normal and contained no denaturation bubbles or other evidence for mismatches. Furthermore, no evidence was found to suggest that our trioxsalen treatment significantly alters the length of the DNA duplex (data not shown). However, the electron microscope cannot resolve either the presence of a few mismatched base pairs or small changes in the length of duplex DNA (less than 130 Å, equivalent to 50 pairs).

During our studies on rRNA R-loops we noticed that the single stranded tails present at the ends of some R-loops appeared preferentially on specific ends of the genes. These tails are presumed to be the product of branch migration since full length R-loops without tails are observed when the R-loops are either fixed with glyoxal or spread in 70% formamide. In 50% formamide spreads, 50% of the 18S rRNA R-loops contained a small RNA knob or tail at the 5' end of the gene, while a tail was seen on the 3' end in only 10% of the R-loops. In addition, under these spreading conditions 33% of the 25S rRNA R-loops were observed to contain RNA tails or knobs at the 3' end, but none were observed on the 5' end.

Finally, the 25S rRNA gene fragment R-loop in pJHC11 comes from the 3' end of the RNA and is only partially stable when spread in 50% formamide. This R-loop corresponds to the same region of the full length 25S rRNA gene which is unstable under these conditions. The paucity of R-loops at distinct ends of the gene suggests that branch migration has some sequence specificity. A comparison of the R-loop map to the denaturation map of yeast rDNA (30) and the  $T_m$ 's of individual restriction fragments (J. Cramer, personal communication) reveals that the regions which branch migrate overlap with regions that have a high A + T base composition. Therefore, R-loops which are high in A + T probably are not very stable in 50% formamide. In agreement with this observation is the instability of mRNA R-loops when spread in 50% formamide (Table 2). The yeast genome is relatively A + T rich (62%) (31,32), so it is not surprising that the R-loops require glyoxal modification for visualization.

It is not clear why branch migration does not go to completion, but

stops at or near a certain point in the rRNA R-loops. This arrest may be due to secondary structure in the unpaired DNA strand since G + C rich DNA would be predicted to contain more stable secondary structure than A + T rich DNA.

This stabilization of R-loops with glyoxal enables them to be manipulated in buffers where they might ordinarily be unstable. This made it possible to separate the DNA containing R-loops from the 100 fold excess of poly(A)<sup>+</sup> RNA driving the hybridization. In addition, spreading from low salt buffers from 50% formamide generally gives better cytochrome:DNA films with improved contrast over that obtained by spreading from a high salt hybridization buffer in 70% formamide.

The effects of the glyoxal treatment were minor causing the appearance of several small denaturation bubbles less than 200 base pairs long. Care should, therefore, be observed when looking at R-loops from small RNA species. In our later studies (to be published elsewhere) the few small bubbles (<200 base pairs) seen when hybridizing total poly(A)<sup>+</sup> RNA to DNA were ignored. These constituted fewer than the equivalent of 2% of the total number of R-loops observed.

The methods detailed in this paper require a minimal amount of manipulation and improve both the ease and effectiveness of R-loop mapping.

### ACKNOWLEDGEMENTS

We are indebted to Fiona Gibson, Lynna Hereford and Michael Rosbash for Poly(A)<sup>+</sup> RNA; the Paul D. Elder Co. for Trioxsalen; and Peter Chandler, James Casey and Pauline Yen for valuable discussions. DBK was supported by a Damon Runyon-Walter Winchell Cancer Research Foundation Fellowship and an U.S.P.H.S. National Research Service Award. This research was supported by grants GM 10991 and GM 20927 from the U.S.P.H.S. and RR 07003, Biomedical Research Grant.

Contribution number 5955 from the Chemical Laboratories at the California Institute of Technology.

\*Present address: Department of Biology, University of Rochester, Rochester, NY 14627, USA

### REFERENCES

1. Thomas, M., White, R.L., and Davis, R.W. (1976) Proc. Nat. Acad. Sci. USA 73, 2294-2298.
2. White, R.L. and Hogness, D.S. (1977) Cell 10, 167-176.
3. Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977) Cell 12, 1-8.

4. Tilghman, S.M., Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V., and Leder, P. (1978) *Proc. Nat. Acad. Sci. USA* 75, 725-729.
5. Early, P.W., Davis, M.M., Kaback, D.B., Davidson, N., and Hood, L. (1979) *Proc. Nat. Acad. Sci. USA*, 76, 857-861.
6. Cole, R.S. (1970) *Biochim. Biophys. Acta* 217, 30-39.
7. Casey, J. and Davidson, N. (1977) *Nucl. Acid Res.* 4, 1539-1552.
8. Blamire, J., Cryer, D.R., Finkelstein, D.B., and Marmur, J. (1972) *J. Mol. Biol.* 67, 11-24.
9. Cramer, J.H., Farrelly, F.W., Barnitz, J.T. and Rownd, R.H. (1977) *Molec. Gen. Genet.* 151, 229-244.
10. Wensink, P.C., Finnegan, D.J., Donelson, J.E. and Hogness, D.S. (1974) *Cell* 3, 315-325.
11. Rubin, G.M. (1975) in *Methods in Cell Biology*, Prescott, D., Ed., vol. 12, pp. 45-64 Academic Press, New York.
12. Heford, L.M. and Rosbash, M. (1977) *Cell* 10, 453-462.
13. Cech, T.R. and Pardue, M.L. (1976) *Proc. Nat. Acad. Sci. USA* 73, 2644-2648.
14. Davis, R.W., Simon, M., and Davidson, N. (1971) in *Meth. in Enzymology*, Grossman, L. and Moldave, K., Eds., Vol. 21, pp. 413-428 Academic Press, New York.
15. Chow, L.T., Roberts, J.M., Lewis, J.B. and Broker, T.R. (1977) *Cell*, 11, 819-836.
16. Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchinson, C.V., III, Slocombe, P.M. and Smith, M. (1977) *Nature* 265, 687-695.
17. Nath, K. and Bollon, A.P. (1976) *Molec. Gen. Genet.* 147, 153-168.
18. Phillippsen, P., Thomas, M., Kramer, R.A., and Davis, R.W. (1978) *J. Mol. Biol.* 123, 387-404.
19. Bell, G.I., Degennaro, L.J., Gelfand, D.H., Bishop, J., Valenzuela, P., and Rutter, W.J. (1977) *J. Biol. Chem.* 252, 8118-8125.
20. Schweizer, E., MacKechnie, C., and Halvorson, H.O. (1969) *J. Mol. Biol.* 40, 261-277.
21. Cramer, J.H., Bhargava, M.M., and Halvorson, H.O. (1972) *J. Mol. Biol.* 71, 11-20.
22. Udem, S.A. and Warner, J.R. (1972) *J. Mol. Biol.* 65, 227-242.
23. Broude, N.E. and Budowsky, E.I. (1971) *Biochim. Biophys. Acta* 254, 380-388.
24. Forscheit, A.B., Davidson, N., and Brown, D.D. (1974) *J. Mol. Biol.* 90, 301-314.
25. Kramer, R.A., Phillippsen, P. and Davis, R.W. (1978) *J. Mol. Biol.* 123, 405-416.
26. Lauer, G.D., Roberts, T.M., and Klotz, L.C. (1977) *J. Mol. Biol.* 114, 507-526.
27. Engel, J.D. and Dodgson, J.B. (1978) *J. Biol. Chem.* 253, 8239-8246.
28. Engel, J.D. and Davidson, N. (1978) *Biochem.* 17, 3883-3888.
29. Shen, J.C. and Hearst, J.E. (1977) *J. Mol. Biol.* 112, 495-507.
30. Cramer, J.H., Farrelly, F.W., Barnitz, J.T., Rownd, R.H. (1977) in *Molecular Approaches to Eucaryotic Genetic Systems*, Wilcox, G., Abelson, J., and Fox, F. Eds., pp. 227-237 Academic Press, New York.
31. Karijord, Ö. and Öyen, T.B. (1975) *Biochim. Biophys. Acta* 383, 255-265.
32. Fukuhara, H., (1967) *Biochim. Biophys. Acta* 134 143-164.